BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

J. Agric. Food Chem. 2001, 49, 5822−5827

Characterization of Peptides Formed during Fermentation of Cocoa Bean

Elif Buyukpamukcu,[†] David M. Goodall,[†] Carl-Erik Hansen,[‡] Brendan J. Keely,*[,†] Sunil Kochhar,[‡] and Hans Wille[§]

Chemistry Department, University of York, Heslington, York YO10 5DD, United Kingdom, Nestlé Research Centre, P.O. Box 44, Vers-chez-les Blanc, CH-1000 Lausanne 26, Switzerland, and Nestlé Product Technology Centre, P.O. Box 204, Haxby Road, York YO91 1XY, United Kingdom

Analysis by SDS-PAGE and GPC-MS of fermented cocoa extracts shows changes in the amount and composition of the major proteins, accompanied by formation of complex distributions of peptides. MS/MS studies and application of SEQUEST sequencing software have allowed identification of two related peptides, a hexapeptide and a nonapeptide, formed from vicilin, one of the cocoa storage proteins. Time course studies of the two peptides show different abundance profiles and indicate, in part, production of the hexapeptide from the nonapeptide.

Keywords: *Cocoa; peptides; GPC-MS; fermentation*

INTRODUCTION

The unique flavor of chocolate is generated during the processing of cocoa beans, the seeds of the tree *Theobroma cacao*, which involves fermentation, drying, and roasting (*1*, *2*). Fermentation is a crucial step in processing, and unfermented beans do not produce cocoa flavor upon roasting (*3*, *4*). Roasting induces high-temperature chemical reactions, believed to be mainly Maillard-type condensations between reducing sugars and the free amino groups of peptides and amino acids, during which flavor compounds are produced (*5*, *6*).

Fermentation involves microbiological and enzymatic reactions that lead to extensive breakdown of the cocoa seed proteins. The peptides so-formed are considered likely to be important flavor precursors (*7*−*9*). Although the role of peptides in the development of cocoa flavor formation has been documented through laboratory experiments involving peptide-containing fractions (*9*−*11*), the structures of peptides formed during fermentation have not been investigated rigorously.

The advent of atmospheric pressure ionization mass spectrometry has greatly facilitated the determination of peptide sequences in protein digest mixtures following chromatographic separation (*12*−*15*). The main approach involves electrospray-mass spectrometry (ES-MS) and tandem mass spectrometry (MS/MS), combined with computer prediction of primary structure through comparison with databases of known protein sequences (*16*−*19*).

Here, we describe the changes in proteins, peptides, and amino acids during fermentation of cocoa beans and the identification of two peptides by gel permeation chromatography (GPC)-MS/MS.

* Corresponding author (telephone +441904432540; fax +441904432516; e-mail bjk1@york.ac.uk).
† University of York.
‡ Nestlé Research Centre.
§ Nestlé Product Technology Centre.

EXPERIMENTAL PROCEDURES

Samples. Amelonado cocoa beans from the Ivory Coast were obtained for each day of a seven-day fermentation period. The beans were hand peeled, milled (5 s) with cooling (IKA M20 laboratory mill, Janke & Kunkel, IKA Werk Staufen, Germany), and sieved (8 mm). Ground cocoa was stored in the dark in airtight containers at 4 °C until use.

Extraction of Proteinaceous Matter from Cocoa. Ground cocoa (10 g) was placed in a cellulose extraction thimble (30 mm × 100 mm; Whatman, Kent, U.K.) and defatted by Soxhlet extraction with *n*-hexane (24 h). The residues were air-dried and stored in the dark in airtight containers at 4 °C.

Defatted cocoa (1 g) was extracted with ethanoic acid/water 1:1 (10 mL) by sonication (Labsonic U, Braun/Inotech, Dottikon, Switzerland; 1 min), homogenized (Polytron, Kinematica AG, Aigle, Switzerland; 1 min), and centrifuged (Beckman Instruments Inc, Fullerton, CA; 27000*g* at 4 °C; 10 min). The supernatant was filtered (0.45 μm; Pall Gelman, Hampshire, U.K.) and stored at −20 °C.

SDS-PAGE Analysis of Proteins and Peptides. Ground samples of cocoa (100 mg) were mixed with extraction buffer [1 mL, 100 mM tris−ethanoate buffer, pH 8.1, containing 1% sodium dodecyl sulfate (SDS) and 1% dithiothreitol]. Mixtures were heated (95 °C for 30 min) to denature the proteins. Samples were allowed to cool and subsequently centrifuged (18000*g* for 5 min). Supernatants were filtered (0.2 μm) and prepared for SDS-PAGE gels by addition of 4 volume equiv of sample buffer (Bio-Rad, glycine system buffer), containing 0.005% (w/v) bromophenol blue as tracking dye. SDS-PAGE gels were prepared and run according to the method of Laemmli (*20*).

Preparation of Standards and Extracts for GPC and GPC-MS. All standards were purchased from Sigma (Sigma-Aldrich Ltd., Poole, U.K.). The L-amino acid standard mixture (comprising 2.5 μmol mL^{-1} of alanine, arginine, aspartic acid, glutamic acid, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tyrosine, and valine and 1.25 μmol mL^{-1} of cystine) was diluted 10 times into mobile phase. Individual peptide standards in mobile phase (diglycine, 1 mg mL^{-1}; triglycine, 0.8 mg mL^{-1}; tetraglycine, 1 mg mL^{-1}; and angiotensin II, 0.1 mg mL^{-1}) were injected (20 μL) onto the GPC column. Cocoa extracts were thawed and filtered prior to injection (20 μL).

10.1021/jf0104127 CCC: $20.00
Published on Web 11/21/2001

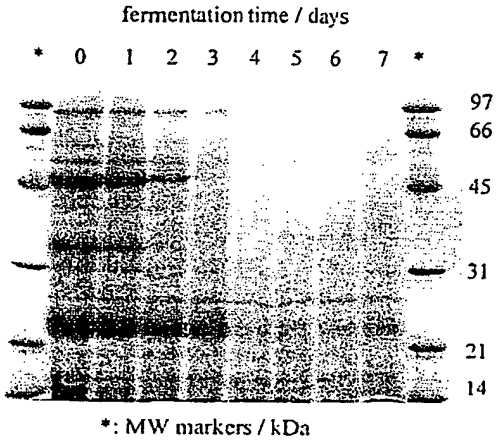

Figure 1. SDS-PAGE of cocoa extracts at different stages of fermentation.

GPC Separation of Amino Acids, Peptides, and Proteins. A Waters (Milford, MA) LC system comprising a 600MS controller, a 717 autosampler, and a 996 photodiode array detector was used for GPC. Instrument control, data processing, and analysis were performed using Waters Millenium 2010 software. All solvents were sparged using helium. GPC was performed using a stainless steel Aquagel-OH 30 column (300 × 7.5 mm; Polymer Laboratories, Shropshire, U.K.) packed with rigid macroporous hydrophilic styrene polymer (particle size = 8 μm, pore size = 30 Å). Isocratic elution was performed at a flow rate of 0.5 mL min^{-1} for 40−60 min using 50 mM ethanoic acid adjusted to pH 3.6 with ammonia as mobile phase. The detector was set to acquire UV−vis absorbance data between 200 and 400 nm.

GPC-MS. GPC-MS was performed on an ion trap mass spectrometer (Finnigan MAT LCQ, San Jose, CA) coupled with a Thermo Separation Products AS3000 autosampler, an SCM membrane degasser, a P4000 gradient pump, and a ThermoQuest UV2000 UV−vis (215 nm). The MS was operated in the electrospray (ES) ionization, positive ion mode using Finnigan Navigator software version 1.2 and LC conditions as described above. ES-MS instrument settings were as follows: sheath gas, 80 (arbitrary units); auxiliary gas, 20 (arbitrary units); spray voltage, 4.2 kV; capillary temperature, 275 °C; capillary voltage, 33 V; tube lens offset voltage, 50 V; collision energy, 45%.

Full-scan MS spectra were acquired over the mass range 50−2000 Da. MS/MS spectra were acquired using data-dependent scanning with an isolation width of 5 m/z. LC-MS/MS data were analyzed using SEQUEST software (BioExplore version 3.1, Finnigan, San Jose, CA) to search against human, cocoa vicilin, and cocoa albumin sequence database subsets created from the OWL database (version 290797) (*21*).

RESULTS AND DISCUSSION

Extracts of unfermented cocoa beans gave four major bands on SDS-PAGE (46, 32, 22, and 14 kDa; Figure 1) as observed previously (*9*, *22*). During fermentation the 46 and 32 kDa proteins, previously identified as vicilins, were hydrolyzed rapidly and were essentially absent after 2 days. The albumin (22 kDa) and the unidentified 14 kDa proteins were not degraded to the same extent, and some traces persisted at day 7 (cf. refs *23* and *24*). The absence of any increase in polypeptides between 46 and 14 kDa over the fermentation period suggests that the proteins break down mainly to components with masses <14 kDa.

Standards. The suitability of the Aquagel-OH column for separation of amino acids, peptides, and proteins was evaluated using amino acid and peptide

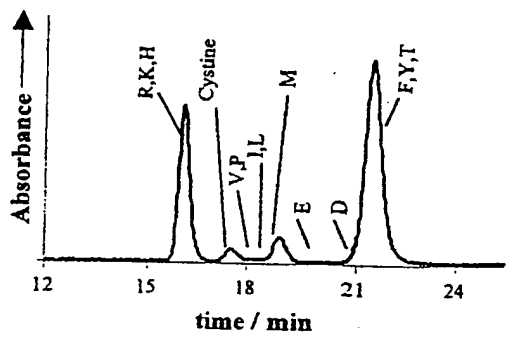

Figure 2. Partial GPC chromatogram at 215 nm of standard amino acids on an Aquagel-OH column.

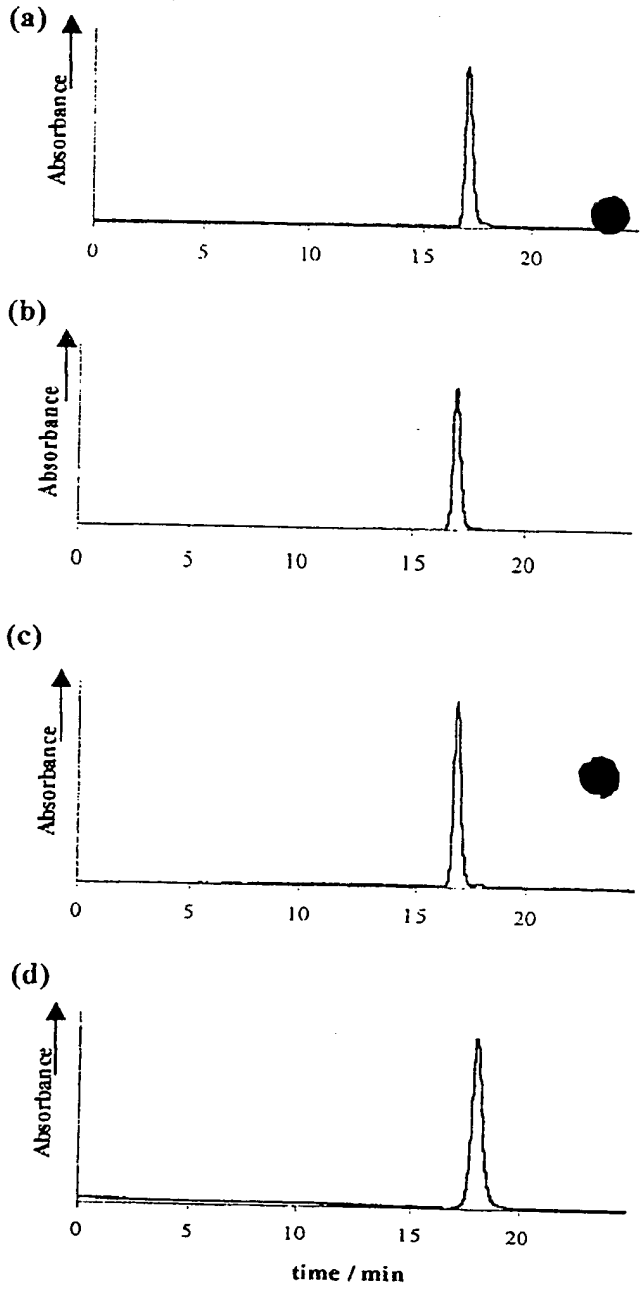

Figure 3. GPC chromatograms of (a) GG, (b) GGG, (c) GGGG, and (d) DRVYIHPF on an Aquagel-OH column. UV detection was at 215 nm.

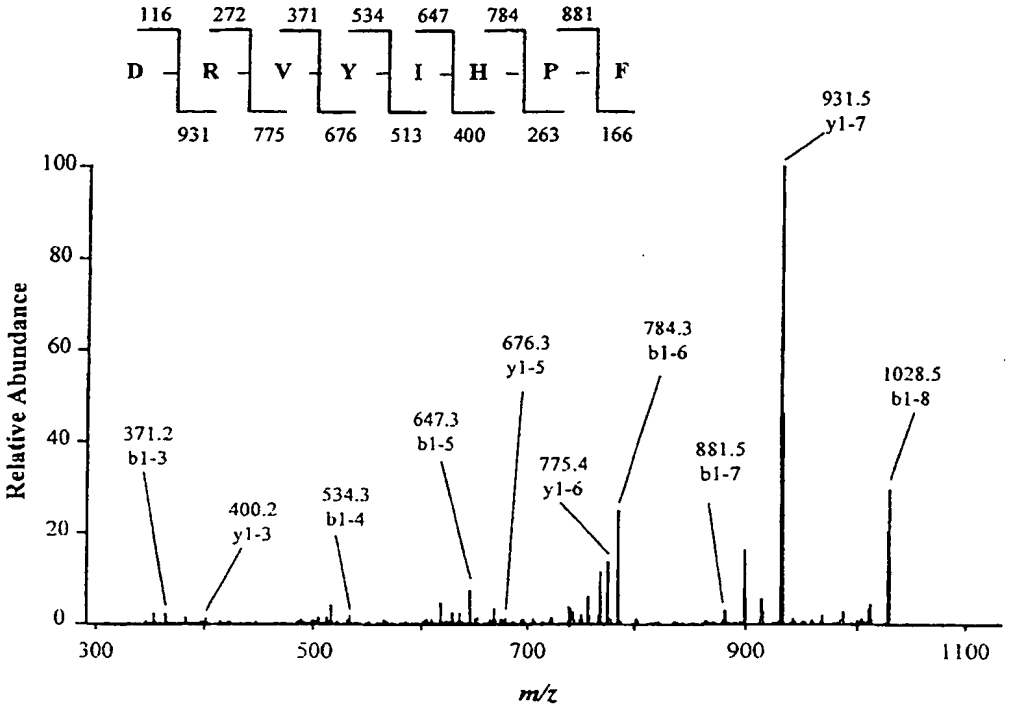

Figure 4. ES-MS/MS spectrum of angiotensin, $[MH]^+ = 1046.5$.

standards covering a range of molecular masses. Aquagel-OH is compatible only with aqueous mobile phase compositions containing up to a maximum of 50% methanol. Initial studies by GPC-MS revealed poor peak shapes with methanol present. The best peak shapes were obtained using a purely aqueous mobile phase buffered with ethanoic acid at pH 3.6 (see Experimental Procedures).

Given the effective nominal exclusion range of the column (100−30000 Da), amino acids would not be expected to exhibit a high degree of separation through a size exclusion mechanism. The standard mixture of 16 amino acids and the dipeptide cystine was chromatographed to examine the effect on retention behavior of the various amino acid side chains. The basic amino acids (R, K, and H), the side chains of which will be protonated at pH 3.6, were the earliest to elute (Figure 2). Amino acids with alkyl side chains (V, P, I, and L) or containing sulfur (cystine and M) eluted next, followed by the amino acids containing a carboxyl (D and E) or hydroxyl functionality (T). The latest eluting amino acids were those possessing an aromatic side chain (Y and F). Although they were present in the standard mixture at the same concentration as the other amino acids, MH^+ ions for A, G, and S were not observed.

Oligomers of glycine (GG, GGG, and GGGG) were all retained to a similar degree (t_R ca. 17 min) but with GG eluting slightly later than GGG and GGGG (Figure 3a−c), revealing low selectivity for GPC in this mass range. Clearly, therefore, for low-mass species, such as the amino acids and glycine oligomers, the column exhibits its most significant activity in the partition chromatography mode. The ability to form $\pi-\pi$ interactions between polystyrene−divinyl benzene and aromatic moieties is believed to be responsible for the particularly long retention of the aromatic amino acids.

The octapeptide angiotensin II (DRVYIHPF) eluted later (t_R = 18 min) than would be expected on the basis of molecular size alone (Figure 3d). The strong retention of angiotensin is influenced by the competing effects of the protonated hydrophilic amino acids R and H and the aromatic amino acids Y and F. Thus, on this GPC stationary phase peptides, as well as amino acids, are separated not only by their size but also by hydrophobic-type interactions, particularly those including aromatic residues.

Peptide sequence information can be obtained by ES-MS using collision-induced fragmentation of the protonated molecules and matching of the MS/MS product ion distributions (*16*, *17*) to b-ions (charge-retaining fragments from the amino terminus) and y-ions (charge-retaining fragments from the carboxy terminus) predicted from known protein sequences (*18*, *19*).

The isolation and fragmentation of a standard of angiotensin II (MH^+ *m/z* 1046) gave a product ion spectrum (MS/MS) arising mainly from cleavage of peptide bonds. Thus, b- and y-ions arising from sequential loss of amino acids from the peptide were present. Analysis using SEQUEST against a human-specific sequence subset extracted from the OWL database accurately predicted the amino acid sequence (Figure 4). The raw correlation score (Xcorr) was 2.03, and the preliminary raw score (Sp) was 369, both of which correspond to good statistical matches and signify a high probability of a correctly identified peptide sequence. The difference (DelCn) between the scores for the top-ranked and the second-ranked peptides was above 0.1, indicating a high probability of a correct match (*16*, *25*).

Cocoa Extracts. Separation of cocoa components by GPC (Figure 5) revealed a complex distribution of components with peaks in the same elution window for the amino acid and peptide standards (t_R values between 12 and 25 min). On-line UV−vis and MS spectra allowed identification of the peaks with high intensity in the UV chromatogram as methylxanthines (theobromine and caffeine, t_R = 27 and 45 min, respectively).

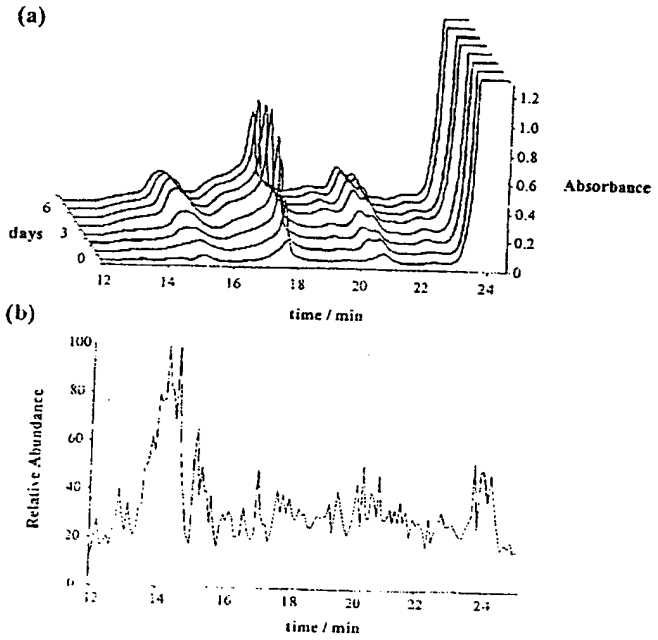

Figure 5. (a) Partial GPC chromatograms (at 215 nm) of cocoa extracts at different stages of fermentation; (b) full scan GPC-MS chromatogram of extract from fermented (5 days) cocoa.

Chromatograms for the different days of fermentation reveal changes in the relative abundances of components in the region where peptides elute (Figure 5a), suggesting that hydrolysis of polypeptides results in the generation of increasing quantities of small peptides and amino acids over the fermentation period.

The base peak mass chromatogram of the fermented extract from day 5 (Figure 5b) shows species that form ions to elute mainly between 12 and 25 min, that is, in the retention time window where the standard amino acids and peptides were found to elute, and reveals greater complexity than is apparent from the UV chromatogram. Two of the most prominent ions were at *m/z* 621 and 902. Both ions show maximum intensity at t_R = 24 min, significantly later than the octapeptide angiotensin II (Figure 3d), suggesting them to be hydrophobic in character and likely to contain aromatic residues. Product ion spectra generated from *m/z* 621 and 902 were analyzed using SEQUEST to extract matches with sequences of cocoa vicilin and albumin proteins. Database subsets for the cocoa proteins were created to facilitate the search. The search results were examined both for close matches between the experimental and predicted monoisotopic molecular masses and for matches between predicted b- and y-ions with those in the MS/MS spectra.

The product ion spectrum (Figure 6a) of *m/z* 902 gave a match to a nonapeptide sequence (APLSPGDVF) in cocoa vicilin protein. SEQUEST search results (fermentation day 5) gave Xcorr = 1.82, Sp = 413, and DelCn = 0.77. The product ion spectrum of *m/z* 621 gave a match to a hexapeptide (SPGDVF) corresponding to a sequence within the nonapeptide (Figure 6b). The best match statistics, obtained for the sample at fermentation day 6, were Xcorr = 1.31, Sp = 258, and DelCn = 0.75. In both cases the match statistics indicate a high probability for the match, with similar or better confidence levels than for angiotensin II. Furthermore, both peptide sequences contain only amino acids with alkyl

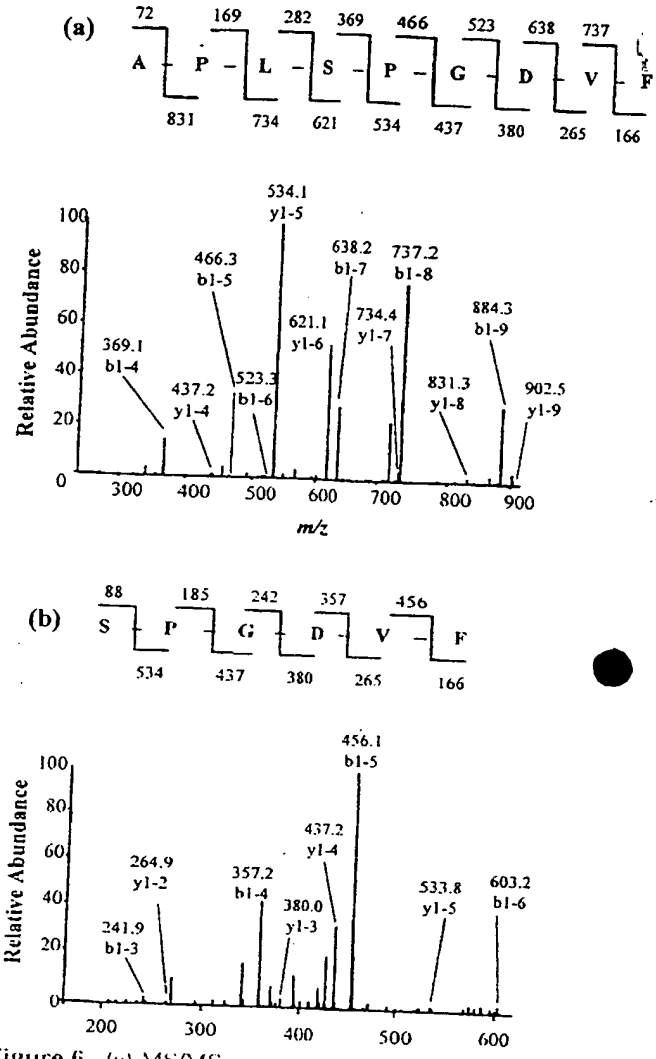

Figure 6. (a) MS/MS spectrum of *m/z* 902.4 illustrating the sequence deduced from cocoa vicilin protein using SEQUEST; (b) MS/MS of *m/z* 621.3 illustrating the sequence deduced from cocoa vicilin protein using SEQUEST.

or hydroxylated side chains and one aromatic residue, consistent with their retention on the GPC column.

It is notable that the two peptides exhibit complete homology, lending even greater confidence to their identifications. The two peptides have been reported previously and were suggested as potential cocoa flavor precursors (*26*). In that study, defatted West African cocoa beans were incubated in vitro with indigenous cocoa enzymes in acidic buffer and the nonapeptide was isolated and characterized by Edman degradation.

The timing of formation of the two peptides (APLSPGDVF and SPGDVF) was evaluated from plots of peak area from the GPC-MS mass chromatograms of *m/z* 621 and 902 (Figure 7) over the seven-day fermentation. The two peptides have very similar maximum abundances but exhibit different formation profiles during fermentation. The profile for the nonapeptide suggests it to be formed early during fermentation and degraded after day three. The initial profile for the hexapeptide is similar, but with an initial lag phase in its formation. The profiles differ markedly from day four onward, with a second increase in abundance of the hexapeptide being apparent. These profiles are more

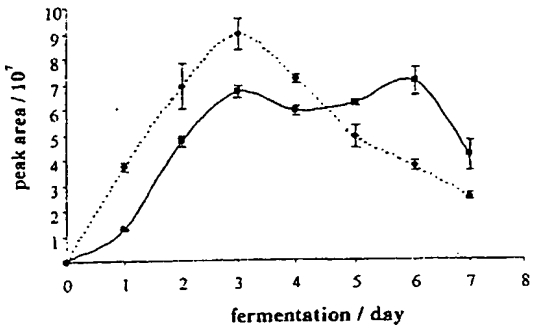

Figure 7. Abundance profiles for two peptides formed during cocoa fermentation: (- - -) *m/z* 902.4; (—) *m/z* 621.3 (error bars represent standard deviation on three replicates).

complex than would be expected by either simple parallel or sequential pathways for formation and decay of the two peptides. Given the homology in the amino acid sequence of the two peptides, their time course study suggests that the hexapeptide is formed from the nonapeptide as the fermentation progresses.

Conclusion. The hyphenated technique of GPC MS has been shown to be extremely powerful for both sequencing and determining changes in abundance levels of peptides in complex mixtures. This paper confirms the identification of two peptide products (a hexapeptide and a nonapeptide) and demonstrates, for the first time, their formation from cocoa vicilin protein during the fermentation step used in chocolate manufacture. Cocoa proteins undergo extensive hydrolysis during fermentation with consequent increases in peptide abundances. Peptide abundance profiles as a function of time indicate that the hexapeptide (which exhibits sequence homology with the nonapeptide) is formed, in part, from the nonapeptide as the fermentation progresses. Future work will be designed to identify the structures of other peptides, follow their time course during fermentation, and clarify in detail the nature of reactions involved in fermentation.

ACKNOWLEDGMENT

We thank Nestlé Research Centre, Lausanne, Switzerland, for provision of cocoa samples and technical support and Trevor Dransfield, University of York, for assistance with mass spectrometry.

LITERATURE CITED

(1) Beckett, S. T. *Industrial Chocolate Manufacture and Use*; Blackie: Glasgow, U.K., 1988.
(2) Baigrie, B. D. Cocoa flavour. In *Understanding Natural Flavours*; Piggott, J. R., Patterson, A., Eds.; Blackie: Glasgow, U.K., 1994; pp 268–282.
(3) Biehl, B.; Brunner, E.; Passern, D.; Quesnel, V. C.; Adomako, D. Acidification, proteolysis and flavor potential in fermenting cocoa beans. *J. Sci. Food Agric.* **1985**, *36*, 583–598.
(4) Ziegleder, G.; Biehl, B. Analysis of cocoa flavour components and flavour precursors. In *Methods in Plant Analysis*; Linskens, H., Jackson, J. F., Eds.; Springer-Verlag: Berlin, Germany, 1988; Vol. 8, pp 320–393.
(5) Hoskin, J.; Dimick, P. S. Role of nonenzymatic browning during the processing of chocolate—a review. *Process Biochem.* **1984**, *19*, 92–104.
(6) Urbanski, J. J. Chocolate flavour/origins and descriptions. the effects of process and bean source. *Manuf. Confect.* **1992**, Nov, 69–82.
(7) Biehl, B.; Wewetzer, C.; Passern, D. Vacuolar (storage) proteins of cocoa seeds and their degradation during germination and fermentation. *J. Sci. Food Agric.* **1982**, *33*, 1291–1304.
(8) Voigt, J.; Biehl, B.; Heinrichs, H.; Kamaruddin, S.; Marsoner, G. G.; Hugi, A. In-vitro formation of cocoa-specific aroma precursors: aroma related peptides generated from cocoa-seed protein by co-operation of an aspartic endoprotease and a carboxypeptidase. *Food Chem.* **1994**, *49*, 173–180.
(9) Voigt, J.; Heinrichs, H.; Voigt, G.; Biehl, B. Cocoa-specific aroma precursors are generated by proteolytic digestion of the vicilin-like globulin of cocoa seeds. *Food Chem.* **1994**, *50*, 177–184.
(10) O'Connor, J.; Warwicker, J. Prediction of peptides in the fermented cocoa storage protein, with reference to the modelled 3D structure. *J. Sci. Food Agric.* **1998**, *77*, 109–11.
(11) Lerceteau, E.; Rogers, J.; Petiard, V.; Crouzillat, D. Evolution of cacao bean proteins during fermentation: a study by two-dimensional electrophoresis. *J. Sci. Food Agric.* **1999**, *79*, 619–625.
(12) Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R. Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 6233–6237.
(13) Dongre, A. K.; Eng, J. K.; Yates, J. R. Emerging tandem-mass-spectrometry techniques for the rapid identification of proteins. *Trends Biotechnol.* **1997**, *15*, 418–425.
(14) Hirayama, K.; Yuji, R.; Yamada, N.; Kato, K.; Arata, K.; Shimada, I. Complete and rapid peptide and glycopeptide mapping of mouse monoclonal antibody by LC/MS/MS using ion trap mass spectrometry. *Anal. Chem.* **1998**, *70*, 2718–2725.
(15) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. Direct analysis of protein complexes using mass spectrometry. *Nat. Biotechnol.* **1999**, *17*, 676–682.
(16) Eng, J. K.; McCormack, A. L.; Yates, J. R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein databases. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976–989.
(17) Yates, J. R.; Eng, J. K.; McCormack, A. L.; Schieltz, D. Method to correlate tandem spectra of modified peptides to amino acid sequences in the protein database. *Anal. Chem.* **1995**, *67*, 1426–1436.
(18) Yates, J. R.; Eng, J. K.; Clauser, K. R.; Burlingame, A. L. Search of sequence databases with uninterpreted high-energy collision-induced dissociation spectra of peptides. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1089–1098.
(19) Adamczyk, M.; Gebler, J. C.; Wu, J. Sequencing of anti-thyroxine monoclonal antibody Fab fragment by ion trap mass spectrometry. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 999–1007.
(20) Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of the bacteriophage T4. *Nature (London)* **1970**, *227*, 680–685.
(21) McCormack, A. L.; Schieltz, D. M.; Goode, B.; Yang, S.; Barnes, G.; Drubin, D.; Yates, J. R. Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femptomole level. *Anal. Chem.* **1997**, *69*, 767–776.
(22) Voigt, J.; Biehl, B.; Wazir, S. M. S. The major seed proteins of *Theobroma cacao* L. *Food Chem.* **1993**, *47*, 145–151.
(23) Amin, I.; Jinap, S.; Jamilah, B. Vicilin-class globulins and their degradation during cocoa fermentation. *Food Chem.* **1997**, *59*, 1–5.

(24) Amin, I.; Jinap, S.; Jamilah, B. Proteolytic activity (aspartic endoproteinase and carboxypeptidase) of cocoa bean during fermentation. *J. Sci. Food Agric.* **1998**, *76*, 123−128.
(25) Yates, J. R. Personal communication, 2000.
(26) Rasmussen, S.; Bach, M. Cocoa flavour precursor peptides, DNA encoding them, processes for producing the peptides, and their use for generating cocoa flavour. WO 96/38472, 1996.

Received for review March 27, 2001. Revised manuscript received August 30, 2001. Accepted October 2, 2001. We thank Nestlé Product Technology Centre, York, for studentship support to E.B. and for partial support toward funding the LCQ and the EPSRC (JREI Grant GR/L26322) for partial support for the LCQ.

JF0104127